Electricity price forecasting accounting for renewable energies: optimal combined forecasts

Carolina García-Martos*, Eduardo Caro and María Jesús Sánchez
Universidad Politécnica de Madrid, Madrid, Spain

Electricity price forecasting is an interesting problem for all the agents involved in electricity market operation. For instance, every profit maximisation strategy is based on the computation of accurate one-day-ahead forecasts, which is why electricity price forecasting has been a growing field of research in recent years. In addition, the increasing concern about environmental issues has led to a high penetration of renewable energies, particularly wind. In some European countries such as Spain, Germany and Denmark, renewable energy is having a deep impact on the local power markets. In this paper, we propose an optimal model from the perspective of forecasting accuracy, and it consists of a combination of several univariate and multivariate time series methods that account for the amount of energy produced with clean energies, particularly wind and hydro, which are the most relevant renewable energy sources in the Iberian Market. This market is used to illustrate the proposed methodology, as it is one of those markets in which wind power production is more relevant in terms of its percentage of the total demand, but of course our method can be applied to any other liberalised power market. As far as our contribution is concerned, first, the methodology proposed by García-Martos *et al* (2007 and 2012) is generalised twofold: we allow the incorporation of wind power production and hydro reservoirs, and we do not impose the restriction of using the same model for 24 h. A computational experiment and a Design of Experiments (DOE) are performed for this purpose. Then, for those hours in which there are two or more models without statistically significant differences in terms of their forecasting accuracy, a combination of forecasts is proposed by weighting the best models (according to the DOE) and minimising the Mean Absolute Percentage Error (MAPE). The MAPE is the most popular accuracy metric for comparing electricity price forecasting models. We construct the combination of forecasts by solving several nonlinear optimisation problems that allow computation of the optimal weights for building the combination of forecasts. The results are obtained by a large computational experiment that entails calculating out-of-sample forecasts for every hour in every day in the period from January 2007 to December 2009. In addition, to reinforce the value of our methodology, we compare our results with those that appear in recent published works in the field. This comparison shows the superiority of our methodology in terms of forecasting accuracy.

Keywords: combined forecasts; design of experiments; electricity price; forecasting; hydro reservoirs; optimisation; time series; wind power

1. Introduction

Since the early 1990s, electricity markets of most of the countries in our economic and social environment have become liberalised power markets in which electricity is exchanged in the same manner as other *commodities*. Nevertheless, electricity presents some particular characteristics that are responsible for its highly volatile and difficult to predict price. Thus, during the last few years an interesting and growing field of research has been electricity price forecasting. Although interest in demand forecasting remains (Cottet and Smith, 2003; Taylor, 2003; Taylor and McSharry, 2007) because it is an important variable in electricity markets, higher interest has been generated by developing methodologies that are able to accurately predict electricity prices. These prices are more difficult to predict, and nowadays the most accurate methodologies for price forecasting clearly obtain larger forecasting errors in comparison with demand forecasting.

Electricity price forecasting is a crucial task both in the short term and in the long run, as it is related to different ways of operating in liberalised markets (Bunn (2004) and Weron (2006) are comprehensive reviews on price forecasting techniques). Although it is not our focus in this paper, another related problem that some authors have dealt with is financial derivatives modelling and the forecasting of energy (Frestad, 2008).

In this paper, we focus not only on short-term forecasting, but also on mid- and long-run forecasting, and the proposed approach is based on time series models. However,

**Correspondence: Carolina García-Martos, Escuela Técnica Superior de Ingenieros Industriales, Universidad Politécnica de Madrid, ETSII-UPM, c/ José Gutiérrez Abascal, 2, Madrid 28006, Spain.*
E-mail: garcia.martos@upm.es

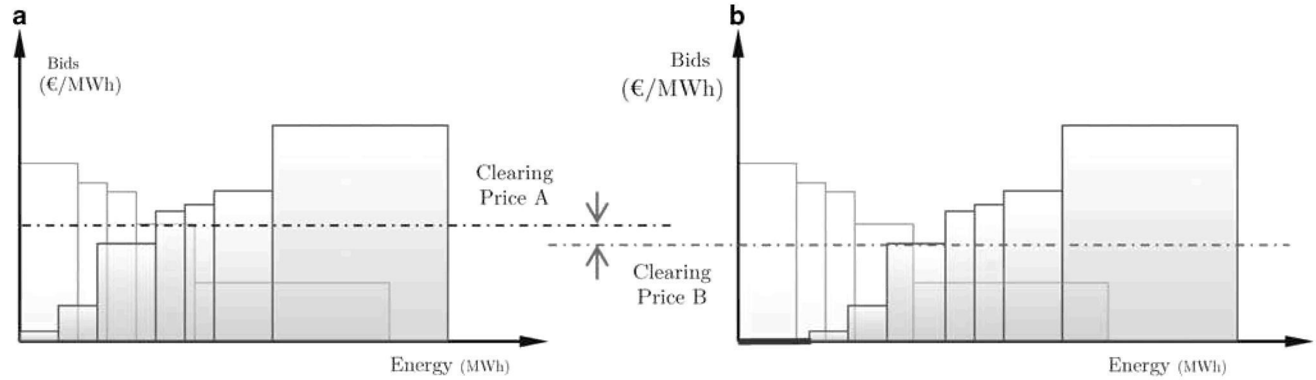

Figure 1 (a) The hourly market clearing price and (b) the expected reduction of the hourly marginal price because of a high wind power penetration.

for short-term electricity price forecasting, alternative approaches are Artificial Neural Networks (ANN) and Support Vector Machine (SVM). Concerning recent works based on ANN techniques, Hippert *et al* (2001) provide a comparative study of some techniques, and Amjady (2006) proposes a forecasting method based on fuzzy logic and ANNs.

Regarding the recent SVM algorithms in Türkay and Demren (2011) a comparison between SVM and ANN methodologies is provided. Recently, Saini *et al* (2010) have applied the SVM model to the national electricity market (Australia). Other artificial intelligence methodologies have also been applied, such as weighted nearest neighbours (Troncoso Lora *et al*, 2007) or swarm-machine approaches (Shrivastava and Panigrahi, 2011).

As far as long-term forecasting is concerned, this issue is linked to the way in which bilateral contracts are negotiated: sellers agree to sell a customer or final consumer a certain amount of energy at a certain price, during a fixed period, which is usually a year. In this context, the disposal of accurate year-ahead electricity price forecasts allows one to reduce the risk that every bilateral contract entails. Nevertheless, and despite the fact that it is a very interesting issue, the literature on long-term electricity price forecasting is scarce, and usually short-term forecasting methodologies do not perform well when the forecasting horizon is extended. Successful attempts at medium-term and year-ahead electricity price forecasting are described by Vehvilainen and Pyykkonen (2005), Conejo *et al* (2010), Alonso *et al* (2011) and García-Martos *et al* (2011).

However, short-term forecasting[1] is linked to *the pool* operation, in which for every hour of the forthcoming day, all the generators and consumers or sellers submit their respective bids. Then, the Market Operator sorts them out, and for the particular case of the Iberian Market the marginal price (the one we focus on and seek to forecast in this paper) is defined by the price submitted by the last generation unit that must be producing for the whole demand to be satisfied (see Figure 1(a)).

The interest in short-term forecasting is related to the improvement of the bids that generators, sellers and users submit to the Market Operator (OMEL in the Iberian Market, on which the empirical applications of this paper focus). The disposal of accurate electricity price forecasts allows the Market Operator to schedule generation units, and this fact is the basis on which every bidding strategy stands (as stated by Conejo *et al*, 2002, 2004). As a result, one-day-ahead electricity price forecasting is a crucial and interesting task. Thus, any reductions achieved in the forecasting error by using new methodological proposals are welcome from the application perspective, and we focus on this issue in this work.

In accordance with the market operation described in Figure 1(a), it should be noted that for every day, there are 24-dimensional data available. Thus, although multivariate modelling is needed (which is not a trivial task), some authors (eg, Huisman *et al* 2007 or Panagiotelis and Smith, 2008) have focused on more conveniently modelling the 24-dimensional vector of prices $(\boldsymbol{y}_d, \boldsymbol{y}_{d+1}, \ldots \boldsymbol{y}_D)$, where $\boldsymbol{y}_d = (y_{1,d}, y_{2,d}, \ldots, y_{24,d})$, instead of the approach that consists of the single/univariate process:

$(y_{1,d}, y_{2,d}, \ldots, y_{24,d}, \ldots, y_{1,D}, y_{2,D}, \ldots, y_{24,D})'$; this process is considered in some other works such as Conejo *et al* (2005) or Contreras *et al* (2003). Other authors have focused on daily modelling the daily prices such as Koopman *et al* (2007).

Furthermore, the increasing concern about climate change and the consequent regulation on greenhouse gas (GHG) emissions, as well as the objective of reducing the dependence on fossil fuels that we do not produce (according to the EU Directive 2009/28/CE, at least 20% of the final energy consumption in 2020 must be generated by renewable energies), is affecting the generation structure in electricity markets. This factor could influence the electricity prices.

Particularly in some European countries such as Spain, Germany and Denmark (Sánchez, 2006; Jonsson *et al*, 2010), which lead the field of renewable energies, there has been

[1]Note that short-term forecasting is our aim in this paper, as wind power is included as an explanatory variable. The extension of the forecasting horizon would only be feasible if wind power forecasting methods were able to compute accurate forecasts in the long run. Such forecasts are not possible at the moment.

a great increase in their installed capacity, which mostly consists of wind farms. This increase has included new variables in the energy markets.

Particularly in Spain, the installed wind power capacity has increased exponentially in the last 15 years: it grew from 407 MW in 1997 to 21 673 MW in 2011 (4950 MW in 2002 and 13 909 MW in 2007). This growth implies that the percentage of the total demand covered by wind farms in Spain increased from 7% in 2005 to more than 20% on some windy days in 2012 (the mean was 17% in 2010). Moreover, given that wind power production pertains to the so-called 'Special Regime', it is required by law that all the electricity produced on Spanish wind farms must be dispatched, and hence, the offers of these producers could be at zero price. This possibility affects the hourly marginal price, by reducing it if the wind power production in a given hour is large and vice versa, as shown in Figure 1(b) (to relate this figure to Figure 1(a), the supply curve moves to the right; thus, the greater the wind power production in hour h of day d, the lower the marginal price).

The second most important clean energy (as a percentage of the coverage of the total demand) in the Iberian Market is hydro production. In fact, hydroelectricity was the premier energy source before the quick increase in wind energy production in the last few years, and its influence on the system operation has traditionally been very important (González *et al*, 2005). Hydroelectricity's percentage of the total demand was approximately 14% in 2010, but it was 18% in 2001. Another market in which the importance of hydroelectric power is even more important is the Nordpool.

In this paper we study the effect of hourly wind power production forecasted by the System Operator (REE), obtained from its forecasting tool SIPREOLICO (available through its website) as well as the percentage of hydro reservoirs, as a measure of the hydroelectric production that could be available if needed to cover peaks in demand and to try to improve the computation of electricity price forecasts. In previous work, Nogales and Conejo (2006) introduced the hourly demand as known for the forthcoming day (in fact, the forecasting error when dealing with demand forecasting is very low, although not zero) and produced one-day-ahead electricity forecasts using transfer function models under this assumption of known hourly demands. More recently, Cruz *et al* (2011) added the hourly wind power production forecasts to improve electricity price forecasting in the Iberian Market.

Neither of these works models the 24-dimensional vector of prices, although recent works have proved its convenience in terms of prediction accuracy both in the short and long run (García-Martos *et al*, 2007, 2011, 2012; Alonso *et al*, 2011). A great drawback of this approach is the high dimension of the problem that should be dealt with, as a large number of parameters should be estimated (the parameter matrices are 24 by 24). Some possible alternatives have been considered by different authors such as ignoring the relationships among different hours of the day (Misiorek *et al* 2006) or modelling the 24 hourly time series separately (García-Martos *et al*, 2007). Panagiotelis and Smith (2008) used a VARX (Vector AutoRegressive with exogenous variables) model specified through a sparse autoregressive coefficient matrix, and a skew t-distributed disturbance has been proposed. More recently, Dynamic Factor Models (DFM) have been extensively used in macroeconomic forecasting (Stock and Watson (2002), among many others), but scarcely applied to energy market forecasting even though they have been revealed as powerful alternatives in the multivariate forecasting of energy issues (Alonso *et al*, 2011; García-Martos *et al*, 2011, 2012). These models are even effective for long-term forecasting, for which most techniques do not exhibit good results.

In this paper, we present several contributions that result in a reduction of the one-day-ahead forecasting errors of electricity prices. First, we extend the methodologies in García-Martos *et al* (2007, 2011) to incorporate exogenous variables both in the Univariate Mixed Model (UMM) and in the multivariate approach given by the DFM. Here, we consider the hourly wind power production and the percentage of the hydro reservoirs. The extensions of the UMM and the DFM, which include the previously mentioned exogenous variables related to wind energy and hydroelectricity, are estimated, and the forecasts are computed for the whole period between 1 January 2007 and the end of December 2009. Then, all the models are compared in terms of their prediction accuracy by means of a Design of Experiments (DOE) (Montgomery, 1984). Moreover, in this paper, the computational experiments performed in García-Martos *et al* (2007, 2012) are extended twofold: first, we do not impose the constraint about fitting the same model for every hour, which could give a sub-optimal solution in terms of prediction accuracy. Second, for those hours in which several models are not significantly different in terms of their accuracy, an optimal combination of forecasts (Bates and Granger, 1969) based on minimising the out-of-sample Mean Absolute Percentage Error (MAPE) is built by extending the ideas in Lam *et al* (2001). The technique of combining forecasts has been used by several authors for computing forecasts of interest in the energy sector, but these forecasts mainly involve load and wind power production (see Taylor and Majithia, 2000; Sánchez, 2006 and 2008). The combination of forecasts for electricity prices obtained by the new procedure introduced in this paper is compared with several previous works:

- García-Martos *et al* (2007, 2012): Although these models do not include exogenous variables, they are well known for their accuracy when applied to the Spanish Market. Indeed, they outperform other previous models such as Nogales *et al* (2002), Contreras *et al* (2003) and Conejo *et al* (2005).
- The recent paper by Cruz *et al* (2011): They provide numerical results for a set of different models. However, the main difference between their methodology and ours is that in addition to wind power production, they incorporate demand as an exogenous variable (particularly in the models with the best performance in terms of their prediction

accuracy). Thus, they need to estimate many more parameters than we do.

The rest of the paper is organised as follows: in Section 2 the data considered (hourly electricity prices, wind power production and hydro reservoirs) are briefly described, and furthermore a short descriptive analysis on them is provided to justify the methodology that is developed. In Section 3, the considered forecasting models are described, that is, the UMM and DFM and its extensions to be able to incorporate exogenous variables are presented. In Section 4, the methodology used to compare all the models, the DOE, is explained. This methodology will be useful for detecting whether the incorporation of wind and hydro significantly improves the results in terms of prediction accuracy, and whether it is necessary to incorporate both the variables or just one to significantly reduce the forecasting error of electricity prices. This analysis will be carried out for each hour, as it could be the case that different models give the smallest forecasting errors in different hours, or at least in groups of hours in which the price and its volatility are similar due to the instantaneous relationship between the price and the demand. Finally, in Section 5, for those hours in which no one model has significantly better forecasting accuracy according to the results in Section 4, a combination of forecasts of several models (namely, those models that give the most accurate results but lack significant differences from each other) is built based on minimising the MAPE. The MAPE is the accuracy metric that is most commonly used in electricity price forecasting (Nogales *et al*, 2002; Nogales and Conejo, 2006). In this section we also provide global accuracy metrics for the model finally built, and we include a comparison with the results of other authors. Finally, Section 6 concludes.

2. The data: hourly electricity prices, wind power production and daily hydro reservoirs

In this section we briefly describe the data we utilise in this paper, as well as some of their main features. This description will be useful to justify the methodology that we will introduce.

The target variable we will forecast with the new methodology we propose is the hourly prices in the period from 1 January 2007 to the end of December 2009 in the Iberian Market. These data are provided in Figure 2(a). In Figure 2(b) and (c), the hourly forecasted wind power production and the daily hydro reservoirs are shown, respectively.

The hourly electricity prices are cleared for day D at 10 a.m. of day $D-1$, when the real hourly wind power productions for each hour of day D are unknown but their forecasts are known. These forecasts are provided by the tool SIPREOLICO (developed by the System Operator REE and freely available through its website), and this tool gives this information to all the agents that are involved in the Market operation, and based on it they submit their bids. We explain the details in Section 3, but the hourly price at hour h and the forecasted wind power for the same hour h have a regular unit root in common. This will be

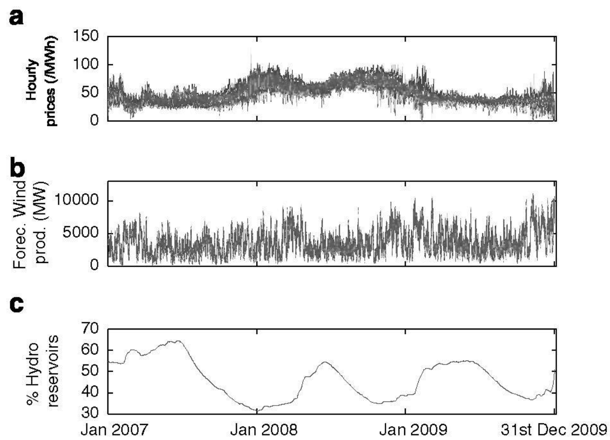

Figure 2 (a) 24 series of hourly prices, Jan 2007–Dec 2009; (b) The hourly forecasted wind power production, given by SIPREOLICO, www.ree.es and (c) the daily hydro reservoirs as percentages of the total capacity.

taken into account to avoid estimating spurious relations between them. In addition, it is possible to consider that the differenced price, $z_{h,t}=p_{h,t}-p_{h,t-1}$, where $p_{h,t}$ is the hourly price at hour h of day t, could be influenced by the differenced exogenous variable at time $t-1$, that is, $x_{h,t-1}=X_{h,t-1}-X_{h,t-2}$, apart from those existing between z_t and x_t. However, we could also reject this idea, as producers and consumers submit their bids for the forthcoming day based on the forecast for that day. But they do not take into account the real value of the previous day. This idea can be checked empirically, as the correlation coefficients between $z_{h,t}$ and $x_{h,t-1}$ are negligible in all the considered hours. Thus, the inclusion of the lagged differenced exogenous variable would not give significant regression coefficients.

For the second variable of interest with regard to improving the electricity price forecast, the hydro reservoirs, the System Operator does not provide its forecast for the forthcoming day. Instead, this forecast will be obtained from the ARIMA model (Auto Regressive Integrated Moving Average) fitted for the historical series, which is updated daily. The selection of the model will be performed by using the Bayesian Information Criteria (BIC; Schwarz, 1978). As occurred in the case of the wind power production, there is a common unit root between the electricity price series and the hydro reservoirs series. Thus, we will proceed with the differenced series to avoid estimating spurious relations between the price and the exogenous variables.

3. Univariate and multivariate forecasting models

In this section, the UMM and the DFM by García-Martos *et al* (2007 and 2012, respectively) are briefly described. To the best of our knowledge, these two models are most likely the most accurate ones for the Iberian Market when exogenous variables are not considered. These models' performances have been demonstrated to be better than the performances of those models that consider the univariate process $(y_{1,d}, y_{2,d}, \ldots, y_{24,d}, \ldots, y_{1,D}, y_{2,D}, \ldots, y_{24,D})'$ with two seasonalities (daily and weekly) as Contreras *et al* (2003), among others, did, instead of considering the 24-dimensional process. Therefore, we propose extensions for these two models that are able to incorporate exogenous variables, particularly the hydro and wind power production. Because the modelling of the 24-dimensional process described by García-Martos *et al* (2007, 2012) was superior to the univariate approach of previous works, we will extend these models to enable them to handle/incorporate the information of the relevant exogenous variables in order to:

1. check whether the incorporation of hydro and wind production significantly improves the forecasting accuracy, and
2. improve other approaches including exogenous variables in terms of their forecasting accuracy. In Section 5 we provide a comparison not only with the best models without exogenous variables but also with a recent work for the Iberian Market (Cruz *et al*, 2011). This model incorporated both the information we consider here and the hourly demand, although not the daily hydro reservoirs.

3.1. The univariate mixed model (UMM)

The UMM proposed in García-Martos *et al* (2007) considers the 24-hourly series of prices: $(\mathbf{p}_d, \mathbf{p}_{d+1}, \ldots \mathbf{p}_D)$, where $\mathbf{p}_d = (p_{1,d}, p_{2,d}, \ldots, p_{24,d})$, and it models each hourly series by a univariate seasonal ARIMA$(p, d, q) \times (P, D, Q)_s$ model, which is also called a SARIMA, whose expression is given by Equation (1(a)). In addition, the optimal historical length used to forecast both for weekdays and weekends was obtained in that paper: 44 weeks. Either complete weeks or weekday data are used as historical data, depending on whether the day for which the forecast is computed is on a weekend or a weekday, respectively. Each hourly series $p_{h,t}$ is modelled as follows:

$$\phi_p(B)\Phi_P(B^s)\nabla^d\nabla_S^D p_{h,t} = \theta_q(B)\Theta_Q(B^s)a_t \qquad (1a)$$

where $\phi_p(B) = (1 - \phi_1 B - \ldots - \phi_p B^p)$, $\Phi_P(B^s) = (1 - \Phi_1 B^s - \ldots - \Phi_P B^{Ps})$, $\theta_q(B) = (1 - \theta_1 B - \ldots - \theta_q B^q)$ and $\Theta_Q(B^s) = (1 - \Theta_1 B^s - \ldots - \Theta_Q B^{Qs})$. Note that a_t is normally independent and identically distributed with zero mean and variance σ_a^2, that is, $a_t \sim NIID(0, \sigma_a^2)$. B is the lag operator and $\nabla = 1 - B$ and $\nabla_s = 1 - B^s$ are the regular and seasonal difference operators, respectively. The usual conditions for these models with respect to stationarity and invertibility apply.

3.2. Accounting for renewable energies (wind and hydro) with the UMM

The UMM can be extended to enable it to account for explanatory variables, that is, by modelling hourly prices using SARIMAX (SARIMA with exogenous variables). In this paper we particularly consider:

- the forecasted wind power production (which is freely available from the System Operator one day in advance) for hour h of day d, which we will refer to from now on as $X_{h,t}$, and
- the hydro reservoirs on day d: H_t.

The empirical findings in the UMM described above for the historical length (44 weeks) as well as the convenience of using complete weeks or weekday data for weekends and weekdays, respectively, also hold in the UMM with the exogenous variables we have introduced.

An interesting and important issue is that $p_{h,t}$, $X_{h,t}$ and H_t share a common regular unit root. Thus, the UMM should be extended to enable it to address this type of exogenous variables (cointegrated ones). This extension is achieved by considering the differenced series of prices, wind power production and hydro reservoirs: $z_{h,t} = \nabla p_{h,t}$, $x_{h,t} = \nabla X_{h,t}$ and $H_t' = \nabla H_t$, respectively. Otherwise spurious regression coefficients between the prices $p_{h,t}$ and $X_{h,t}$ and H_t would be estimated.

Four different SARIMAX models are estimated, and in Section 3 we compare them to discover which is the best alternative (if there is one) with respect to its forecasting accuracy. In the first model the forecast for the wind power production for hour h and day t is the first regressor, and in the second model the first regressor is the forecast for the total daily wind power production, which is computed as a sum of the hourly forecasts, that is, $Y_t = \Sigma_{h=1}^{24} X_{h,t}$. In both cases, the hydro reservoirs is the second regressor (be aware that the hydro reservoirs is a daily series). The first and second UMM models with explanatory variables would have the following expressions for each differenced series of hourly prices, $\nabla p_{h,t} = z_{h,t}$:

$$\phi_p(B)\Phi_P(B^s)\nabla_s^D \tilde{z}_{h,t} = \theta_q(B)\Theta_Q(B^s)a_t + \beta_{Wind,h} x_{h,t} + \beta_{Hydro} H_t' \qquad (2)$$

$$\phi_p(B)\Phi_P(B^s)\nabla_s^D \tilde{z}_{h,t} = \theta_q(B)\Theta_Q(B^s)a_t + \beta_{SumWind} y_t + \beta_{Hydro} H_t', \qquad (3)$$

where $y_t = \nabla Y_t$, and $\beta_{Wind,h}$, β_{Hydro} and $\beta_{SumWind}$ are the coefficients of the corresponding regressors according to (2) and (3). The autorregressive and moving average components are as in subsection 3.1.

The third and fourth models correspond to including only the hourly wind power and the sum of the daily wind power, that is, Equations (2)–(3), respectively, without the hydro term.

For all these models, both the panel $\tilde{z}_t = (\tilde{z}_{1,t}, \tilde{z}_{2,t}, \ldots, \tilde{z}_{24,t})$ and the panel of forecasted hourly wind power production $X_h = (X_{1,t}, X_{2,t}, \ldots, X_{24,t})$ comprise daily data. Their respective hth column is the centred differenced hourly prices for hour h of day t: $\tilde{z}_{h,t}$, and the forecasted wind power production for hour h of day t: $X_{h,t}$.

The frequency of the series that includes the hydro reservoirs, H_t, is daily, as it is the corresponding differenced series H_t'.

3.3. The dynamic factor model for seasonal data (DFM)

Although the UMM and (even more) its extension can incorporate regression effects that are very accurate approaches for the one-day-ahead forecasting of electricity prices, they do not account for the multivariate cross-correlation structure of the panel of 24 hourly series of prices. Thus, García-Martos *et al* (2012) proposed the DFM for seasonal data, which can be roughly summarised as the dynamic extension of the static Principal Component Analysis. The DFM can also be described as the seasonal extension of the Lee-Carter (1992) or Peña-Box (1987) model.

The main advantage of the seasonal DFM is that it allows one to account for the multivariate dynamic structure but avoid the estimation of a Vector-ARIMA (VARIMA) model, which in the case of interest (a 24-dimensional vector of hourly series) would imply that each parameter is a 24 by 24 matrix. This would lead to the well-known problem of 'curse of dimensionality', which is not feasible from the applied perspective. Hence, it is necessary to apply a dimensionality reduction technique such as the Dynamic Factor Analysis.

The methodological details of the seasonal DFM without regression variables can be seen in García-Martos *et al* (2012). In the next subsection we include the equations of the extended model, which incorporates regression variables even if they are non-stationary as electricity prices are.

3.4. DFM with explanatory variables

The equations of the model that considers the 24 differenced hourly wind power productions $x_{h,t}$, $h=1, \dots, 24$, as well as the differenced hydro reservoirs, $H_t^{'}$, as regressors are as follows:

$$\tilde{\mathbf{p}}_t = \mathbf{f}_t \mathbf{P}^{\mathrm{T}} + \mathbf{e}_t,$$

$$\phi_p(B)\Phi_P(B^s)\nabla_s^D \tilde{f}_{K,t}^{'} = \theta_q(B)\Theta_Q(B^s)a_t + \beta_{Wind,h}x_{h,t} + \beta_{Hydro}H_t^{'} \ \forall k = 1, \dots, r.$$

where $\tilde{\mathbf{p}}_t$ are the centred prices, $\mathbf{f}_t$ is the vector of unobserved common factors, r is the number of common factors, $\tilde{f}_{k,t}^{'}$ refers to each centred differenced common factor and $\mathbf{P}^{\mathrm{T}}$ is the transpose of the loading matrix whose columns are the weights of the original series that are needed to build the unobserved common factors $f_{k,t}$. The autoregressive and moving average components are as in subsection 3.1. When the original series are non-stationary the common factors are also non-stationary. Then, the SARIMAX model for the common factors should be estimated for the differenced common factors as well as for the differenced regressors to avoid the estimation of spurious relations between variables, which is also considered in the UMM with explanatory variables. The $\mathbf{e}_t$ are stationary processes that explain a small proportion of the variability of the vector of hourly prices $\mathbf{p}_t$, and these $\mathbf{e}_t$ are modelled by univariate stationary ARMA processes.

It is worth mentioning that the r common factors at time t, $f_{k,t}$, $k = 1, \dots, r$, are built by weighting with the corresponding loads the original hourly series for day t with the corresponding loads. Hence, the common factors are daily series. As discussed in the previous subsection, the information about the forecasted wind power production can be taken into account in two different ways: either with 24 regressors that are the forecasted hourly wind power productions $X_{h,t}$ (in fact the differenced series $x_{h,t}$ or with the total forecasted wind production for the whole day: Y_t (and the differenced process $\nabla Y_t = y_t$). Moreover, the hydro reservoirs will be considered but not as in the UMM specification.

4. Design of experiments

The basic idea of factorial designs is to obtain values of the response variable for all the possible combinations of the levels of the factors under study. Each combination is generally named a 'treatment', and in our particular problem we will refer to these combinations as the 'forecasting method'.[2]

In the previous section, we presented different univariate and multivariate models that are used for the one-day-ahead forecasting of electricity prices. Some of these models consider information concerning renewable energies and others do not. Our main objective is to determine both:

- which Model is better (UMM or seasonal DFM either with two or three common factors) and
- whether to consider the hydro reservoirs as well as the forecasted wind power production.

Our goal is to obtain the smallest forecasting errors in the prices.

For this purpose, a computational experiment has been conducted in which we computed one-day-ahead out-of-sample forecasts for every hour in the period from the beginning of 2007 to the end of 2009. In this experiment, we focused on the accuracy metrics that would be of interest according to previous works in this field (Nogales and Conejo, 2006). In this sense, the hourly accuracy metrics we consider are the percentage errors $e_{h,t}$, which are used to compute the daily MAPE as follows:

$$MAPE_t = \frac{1}{24}\sum_{h=1}^{24} e_{h,t}, \text{ where } e_{h,t} = \frac{|\hat{p}_{h,t} - p_{h,t}|}{p_{h,t}}. \quad (4)$$

$\mathrm{MAPE2}_t$ is the daily median of the $e_{h,t}$ on day t. One of the novelties of this paper is that we build the optimal model without assuming that the same 'Forecasting Method' is necessarily the best one for all the hours (these behaviour patterns in terms of level and variability are different due to the instantaneous relationship between the load and the prices).

Thus, for every hour we analyse the influence of several factors in the hourly forecasting error $e_{h,t}$, instead of just the MAPE we obtained when we used the same forecasting model and exogenous variables for every hour. The factors whose influence we analyse are mentioned above. Here we also make explicit the levels we considered for them, which are as follows:

[2]A forecasting method will be defined by a Model (UMM, DFM with two common factors or DFM with three common factors). We consider the following cases: no wind power production, hourly production or daily information on this production, and considering: no hydro reservoirs data or doing so.

- The Model (UMM and DFM with both $r=2$ and 3 common factors),
- Wind power production (not considering this information, considering the hourly forecasted wind power production, and considering the total forecast for the production of the forthcoming day),
- Hydro reservoirs (both accounting for this information and not doing so).

The main aim is to select the appropriate level for: (a) the Model, (b) the incorporation (and how, hourly or total daily) or not of the forecasted wind power production and (c) the incorporation or not of the hydro reservoirs data that minimise the forecasting error. As opposed to previous works, we do not impose the constraint of having the same combination of Model, Wind and Hydro Regressors for all the hours (24), since this constraint could lead us to a sub-optimal solution.

Thus, the out-of-sample forecasts as well as forecasting errors $e_{h,d}$ for all the hours $h=1,\ldots,24$ in the period under study (every day in the years 2007–2009, $d=1,\ldots T$, where T is the number of days for which a forecast is computed) are calculated with the $3\times3\times2=18$ forecasting methods that correspond to three levels of the factor 'Model', three levels of the factor 'Wind' and two levels of the factor 'Hydro' aforementioned, respectively. Then, all these models are compared by developing an Analysis of Variance (ANOVA) (see Montgomery (1984) for a detailed explanation), which allows us to conduct comparisons that are as homogeneous as possible. In addition, this comparison is able to detect which factors influence the response variable, which is the hourly percentage error $e_{h,t}$ in our case.

The equation of the linear model to be estimated for each hourly forecasting error is as follows:

$$e^h_{ijkd} = \mu+\alpha_i+\beta_j+\gamma_k+\delta_d+(\alpha\beta)_{ij}+u_{ikjd}, h = 1, \ldots ,24. \quad (5)$$

$$u_{ikjd} \sim NIID(0,\sigma^2),$$

where e^h_{ijkd} is the percentage error in hour h when forecasting with 'Model' i not including wind production forecasts, including hourly or daily forecasts: ($j=1,2,3$), and including and not including ($k=1,2$) the hydro reservoirs data on day d.

The parameter μ is the grand mean, and $\alpha_i, \beta_j, \gamma_k$ and δ_d are the main effects related to the factors 'Model', 'Wind' and 'Hydro' and the block 'Day',[3] respectively. In addition, $(\alpha\beta)_{ij}$ refers to the interaction between the factors 'Model' and 'Wind'.[4] The following restrictions apply when estimating the model:

$$\sum_{i=1}^{3}\alpha_i=\sum_{j=1}^{3}\beta_j=\sum_{k=1}^{2}\gamma_k=\sum_{d=1}^{T}\delta_d=\sum_{i=1}^{3}(\alpha\beta)_{ij}=\sum_{i=1}^{3}(\alpha\beta)_{ij}=0.$$

[3]As in previous works, the effect of the Day is included as a block.

[4]Models with other interactions were also considered but they were not significant. The F-statistics were smaller than one. Hence, they could be removed from the model and considered negligible.

Thus, α_i measures the increase/decrease in the average response of Model i with respect to the average level. The same interpretation holds for the other main effects, β_j and γ_k, which are due to factors 'Wind' and 'Hydro', respectively. In this work, the objective is to develop an optimal model that obtains the preferred level for 'Model', 'Wind' and 'Hydro'. However, the 'Day' may affect the prediction error. For instance, if there is an unexpected price on one of these days, none of the forecasting methods will provide an accurate forecast for the corresponding hour. For the purpose of eliminating the effect of Day, we include this variable as a block. This let us both:

- Not increase the residual variability that is due to the 'Day', being able to study the influence of the factors 'Model', 'Wind' and 'Hydro', and
- Make the selection of the most appropriate forecasting method more robust, because it was not affected by spikes in forecasting errors caused by spikes in prices. Even in these cases, the procedure proposed here performs well in terms of prediction accuracy, as the forecasts have been computed for every day and hour in the large span of years under study in this paper, including those days on which the price could be considered as outliers.

The error term u_{ikjd} includes all the effects that are not explicitly included but can somehow explain some of the variability of the response variable e^h_{ijkd}. We assume that the errors are independent, Gaussian and homoscedastic. Of course the ANOVA is performed for all 24 h of the day. We analyse the influence of the aforementioned factors (Model, Wind and Hydro) in the hourly forecasting error, e^h_{ijkd}, $h=1,\ldots,24$, and hence 24 ANOVAs are performed. If we only studied the influence of these factors on the MAPE$_d$, as in previous works, we would be imposing the constraint of using the same combination of levels for all hours, which is unnecessary and could lead to worse results (as in García-Martos *et al*, 2007, 2012).

Once the model in Equation (5) has been estimated by Maximum Likelihood (ML) and the diagnostics have been checked (residuals must be Gaussian, homoscedastic and independent), the results can be interpreted and the best forecasting model selected (when selecting the most appropriate levels of the factors under study, namely, Model, Wind and Hydro) whenever it exists. In some cases several forecasting methods (where a forecasting method considers a particular level for all the factors under study, which is generally called a 'treatment' in DOE) among the 18 considered could be non-significantly different. In these cases we will combine all the best forecasts, as explained in the next section.

Given that 24 ANOVAs have been performed, the results of only some of them are shown (they were selected according to the representativeness of the corresponding hours).

For instance, all the hours in the deep night (3–7) exhibit similar behaviour and similar conclusions can be stated not only in terms of the significant factors in the DOE ('Model' and 'Wind') but also because the pattern of the interaction plot is similar for all these hours. The factors 'Model' and 'Wind' as

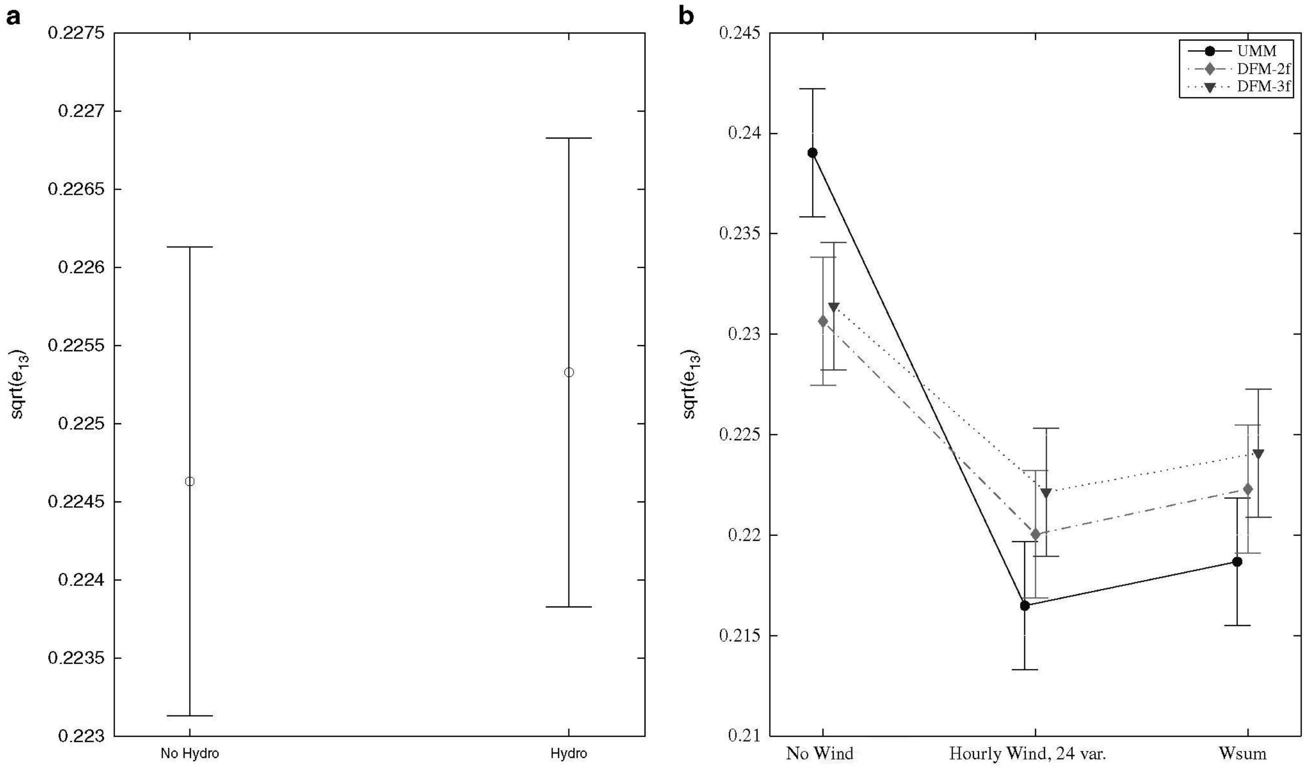

Figure 3 (a) The main effect due to the factor 'Hydro' and (b) the interaction plot corresponding to the significant interaction Model—Wind in hour 3.

well as their interaction are significant (the respective *p*-values are 0.0016, 0 and 0.0055, which are smaller than the most used significance level 0.05, which corresponds to a confidence level of 95%). According to this result, in Figure 3 the results for the ANOVA of hour 3 are shown.[5] Given that the interaction between 'Model' and 'Wind' is significant, the preferred levels for both of them must be selected in the interaction plot (Figure 3(b)) and this selection must be made jointly for both factors. Given that all the other interactions were not significant and thus were not considered in the estimation and do not appear in linear model 5, the main effect 'Hydro' (Figure 3(a)) can be interpreted and used to determine the most appropriate level of this factor. In terms of the mean, the smallest percentage error is given by the forecasting method that does not include hydro and incorporates wind by means of the hourly information when it uses the DFM with three unobserved common factors. Still there are no statistically significant differences between this forecasting method and many others. In fact, we can discard all the UMM models without wind both in the case of incorporating hydro and not doing so, as they are significantly worse. This result is shown in the interaction plot, where the UMM without wind does not overlap with any other one. In the next step of our methodological proposal, the models that are significantly worse are not considered for the combination.

For illustration purposes, the results of the ANOVA obtained for hour 14, which is another representative one, are shown in Figure 4. These results are representative of those results that were obtained for hours 13–15, which is sensible bearing in mind the shape/profile of the load curve. The factors 'Model' and 'Wind' of the DOE are significant as well as their interaction (the *p*-values are equal to 0 in all cases), whereas the factor 'Hydro' is not significant (the *p*-value is larger than 0.05). Moreover, according to Figure 4(b), the forecasting method that leads to the smallest prediction errors (average) for this hour is the UMM that incorporates hourly forecasts for the wind power production without using the hydro reservoirs data. Nevertheless, this method is not significantly better than the others, given the overlapping of the observed intervals. However, Figure 4(b) is very useful for discarding those forecasting methods that are significantly worse than others. (The list of inferior methods consists of all the methods that do not include any information about wind power forecasting, ie, all the methods previously proposed in García-Martos *et al*, 2007, 2012. These methods are in fact six forecasting methods out of the 18 considered in this paper.) The rest (12 methods) will be adequately combined in the manner explained in the forthcoming section, by computing the adequate weights via optimisation.

[5]Note that the square root of the percentage error is analysed. First, the percentage error was considered, but the diagnostic checking was not correct because the residuals were not homoscedastic. The proposed transformation solves this problem.

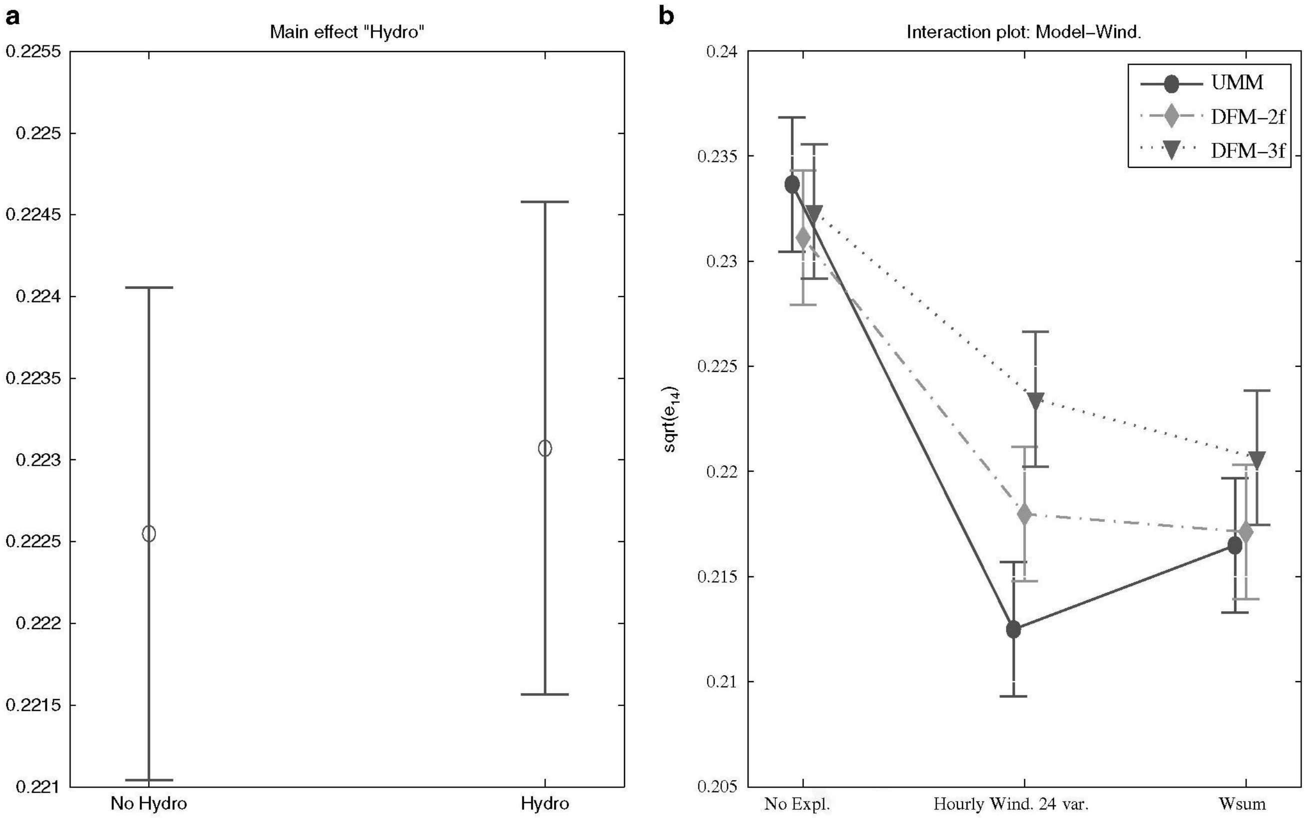

Figure 4 (a) The main effect due to factor 'Hydro' and (b) the interaction plot corresponding to the significant interaction Model—Wind in hour 14.

An ANOVA like the two shown for hours 3 and 14 is carried out for every hour, and the significantly better models are obtained whenever they exist. Similarly, the significantly worse models are discarded from the forecast combination. Similar patterns are detected for hours in which the behaviour (and load) is similar. For instance, for hours 24, 1 and 2 (the first hours of the night), just two forecasting methods (out of 18), namely, UMM with hourly wind power information, and with and without hydro, are combined because they are significantly better than any other method.

In this sense, it is very interesting to see how for similar hours (in terms of load) the same models are significantly worse or better than others. We obtain different clusters (groups) that can be interpreted in terms of the instantaneous relationship between the load and the price.

In the next section these best models obtained for each hour by means of the corresponding 24 DOEs are combined to obtain an optimal model in terms of forecasting accuracy.

5. Minimising the MAPE in combined forecasts

Combinations of forecasts have traditionally been used to weight the information of several models instead of selecting a single one and discarding the rest, and this practice was the key motivation for this issue in the seminal paper by Bates and Granger (1969). Since then, forecast combinations have been used in different applications and contexts (see Clemen (1989), De Menezes *et al* (2000), Stock and Watson (2004), among many others). Furthermore, some developments in the power markets context have occurred, and they are mainly focused on load and wind power forecasting (see Taylor and Majithia, 2000; Sánchez, 2006 and 2008). In this paper, as a novelty, a combination of forecasts is used in the electricity price context (where it is not usual) to combine those models that are not significantly different in terms of their prediction accuracy, according to the large computational experiment performed in a previous step. This combination makes the weights robust and more stable over time; as in the previous step to compare all the forecasting methods, the effect of the day was eliminated by including it as a block, as explained previously. Moreover, the DOE was conducted for a long period, which allows us to use the results in subsequent years.

Once the groups of the best forecasting methods have been obtained for each hour by means of the computational experiment presented above, the forecasts that these methods produce are going to be combined to minimise the MAPE. The MAPE is the most extended accuracy metric when checking the accuracy of electricity price forecasts obtained with new methodological proposals and comparing different forecasting techniques (Conejo *et al*, 2005; Nogales and Conejo, 2006).

Out-of-sample one-day-ahead forecasts have been computed for every hour in the period under study (ie, the years from 2007 to 2009) for the 18 forecasting methods given by the two Models, the three levels of the Factor 'Wind' and the two levels for the Factor 'Hydro' that are explained in detail in Section 4. After a careful analysis of the ANOVAs, it can be stated that some models are significantly better than others (depending on the hour). When there is not a single best model, all these 'best methods' should be combined to improve the forecasting accuracy. An additional novelty in our approach is that we take advantage of the DOE results, where the inclusion of 'Day' as a block allows us to make the combination of 'the best models' robust.

According to the expression of the MAPE for day t given in (4), it is computed as the mean of each $e_{h,t}$, h=1, … ,24. Thus, because each $e_{h,t}$ is an absolute value, we should minimise each $e_{h,t}$. In the most general case in which none of the forecasting methods is discarded, we have 18 forecasts available for each $p_{h,t}$. A combination of these forecasts would be given by:

$$\hat{p}_{h,t}^{comb} = \omega_{h,1}\hat{p}_{h,t}^{m_1} + \cdots + \omega_{h,18}\hat{p}_{h,t}^{m_{18}}, \omega_{h,1} + \cdots + \omega_{h,18} = 1,$$

where $\hat{p}_{h,t}^{m_M}$ is the forecast given by the Mth model for the hour h of day t and $\omega_{h,M}$ the weight of the forecast of this Mth model.

To obtain the weights to build the combination of forecasts that minimises the MAPE, we proceed by extending the approach suggested by Lam *et al* (2001), to minimise:

$$e_{h,t}^{comb} = \frac{\left|\sum_{M=1}^{18} \omega_{h,M}\hat{p}_{h,t}^{m_M} - p_{h,t}\right|}{p_{h,t}}, h = 1, \ldots, 24.$$

The resulting non-linear optimisation problem corresponding to the hth hour is as follows:

$$\underset{w_{h,1},\ldots,w_{h,18}}{minimize} \frac{\left|\sum_{M=1}^{18} \omega_{h,M}\hat{p}_{h,t}^{m_M} - p_{h,t}\right|}{p_{h,t}}. \tag{6}$$

Thus, this approach is based on the solution of 24 non-linear mathematical programming problems. The objective function is non-linear (because it involves the absolute value function), but it can be linearised straightforwardly by including an additional optimisation variable d_h:

$$\underset{w_{h,1},\ldots,w_{h,18},d_h}{minimise} \; d_h \tag{7a}$$

subject to:

$$-d_h \leq \frac{\sum_{M=1}^{18} \omega_{h,M}\hat{p}_{h,t}^{m_M} - p_{h,t}}{p_{h,t}} \leq d_h. \tag{7b}$$

The following observations are given in order:

- The resulting formulation corresponds to a linear mathematical programming problem that can be efficiently solved using any commercial optimisation software (such as GAMS, AMPL or MATLAB).
- Optimally, $d_h = e_{h,t}^{comb}$ for every hour.
- The number of continuous optimisation variables depends on the hour under consideration.

Therefore, the idea proposed here consists of the following steps:

1. Implementing 18 different univariate and multivariate forecasting methods[6] and computing their corresponding forecasts for every hour in a large span of days (from the beginning of 2007 until the end of 2009),
2. The results of the computational experiment are analysed using ANOVA, so as to obtain those models that are statistically worse in terms of their forecasting accuracy (to discard them in the forecasts combination) or if there is a model that is significantly better than the others use it directly to forecast.
3. Build a combination of forecasts for each hour by appropriately weighting those models that are significantly better than the discarded ones according to the ANOVA. This process is conducted by estimating the corresponding weights for the forecast combination by minimising the daily MAPE, which is the accuracy metric whose use is the most extensive in this field. Thus, to build the forecasts of the forthcoming day, an 18 by 24 matrix of weights for the forecasts is obtained by optimisation as described above, although some of the elements of this matrix are zero according to the DOE of the corresponding hour (the worst models are discarded and only the best ones are combined).

6. Numerical results

In this section, we present some results for the period under study and compare our results with previous works.[7] Although we do not incorporate the load, we will show the better performance of our approach in terms of the prediction errors.

In Table 1, a comparison with the monthly forecasting errors obtained in previous accurate works for the Iberian Market is provided, and we focus on the UMM (García-Martos *et al*, 2007) and the DFM (García-Martos *et al*, 2012). The months shown are those that appear in these previous works, which are the ones with the best performance for the Spanish Market in the short run, and this is the reason for using them as benchmarks. For all the months considered except

[6]These methods come from the three levels of 'Model' (UMM, DFM with two unobserved dynamic factors and DFM with three unobserved dynamic factors), three levels of 'Wind' (no wind power forecasts included, hourly forecasts of wind power production and daily forecasts) and two levels of 'Hydro' (not including the reservoir level as exogenous variable and including it).

[7]We use García-Martos *et al.* (2007, 2012), as these are the most accurate models for short-term forecasting in the Iberian Market, and their accuracy has been demonstrated not just for a few days or weeks but for large spans of years under different circumstances. Another benchmarking model is the very recent work by Cruz *et al* (2011), where they use several models and take into account both the wind power production and the load forecasts.

Table 1 The monthly MAPE (a comparison between combined forecasts and previous works)

Month	*Nov* 2007	*Dec* 2007	*Jan* 2008	*Feb* 2008	*Mar* 2008	*Apr* 2008	*May* 2008
DFM (García-Martos *et al*, 2012)	8.815	11.034	9.459	7.691	7.992	6.440	5.605
UMM (García-Martos *et al*, 2007)	9.331	11.260	10.090	7.607	9.032	6.937	6.176
Comb. forecasts	7.76	11.2	8.97	7.03	6.69	5.84	5.53
Month	*Jun* 2008	*Jul* 2008	*Aug* 2008	*Sept* 2008	*Oct* 2008	*Nov* 2008	*Dec* 2008
DFM (García-Martos *et al*, 2012)	5.506	4.739	5.444	5.395	6.753	7.112	11.231
UMM (García-Martos *et al*, 2007)	5.867	5.336	5.711	5.501	6.902	7.546	11.962
Comb. forecasts	4.90	4.40	5.00	4.36	5.77	5.96	10.48

Table 2 Monthly MAPEs (another comparison between results obtained with the new approach proposed and Cruz *et al*, 2011)

Month	*Nov* 2007	*Dec* 2007	*Jan* 2008	*Feb* 2008	*Mar* 2008
Cruz *et al* (2011)	8.19–8.43–9.07	9.10–9.18–10.63	7.35–7.87–9.13	6.30–6.75–7.49	7.36–7.63–7.69
Comb. forecasts	7.76	11.2	8.97	7.03	6.69
Month	*Apr* 2008	*May* 2008	*Jun* 2008	*Jul* 2008	
Cruz *et al* (2011)	6.02–6.29–6.33	5.82–5.90–6.11	5.27–5.29–5.31	4.83–4.88–4.90	
Comb. forecasts	5.84	5.53	4.90	4.40	

December 2007,[8] the combined forecasts outperformed the UMM and DFM. Table A1 provides additional information concerning this comparison.

Thus, considering the span of months included in Table 1, the new methodology of computing combined forecasts provided in this paper is clearly superior to the previous approaches, as it only gives the smallest average MAPE for the period considered, but the differences are also statistically significant from the alternative approaches (the UMM and the DFM).

Furthermore, the average MAPE and the standard deviation are both smaller, although detailed monthly results are not included here for the latter statistic.

The forecasting results corresponding to the rest of the period under study here (in 2009) that were not included in previous works and were not compared with the UMM and the DFM in Table 1 appear below in this section.

Another interesting comparison is made against the set of predictions given in Cruz *et al* (2011). The coincident period in time between their work and ours comprises the months from November until July 2008. In Table 2, we include our monthly forecasting errors, which we obtained when we implemented the new methodological proposal presented in this paper for the period under consideration. This table also includes the forecasting errors of Cruz *et al* (2011) for the three best models they consider. (Note that they give results for six models. We provide in this table the results of their best three models in each month, since the winner is not always the same one, and even the three best ones do not always correspond to the same models.) It is important to emphasise that in addition to wind power forecasts, they include demand forecasts, which requires estimating a much larger number of parameters (regression coefficients). Moreover, given the inclusion of the demand, the comparison could be unfair to our novel approach, but our model still works better in the vast majority of cases even though the period of the comparison was imposed by the existing results in Cruz *et al* (2011) and not selected by us. The comparison between our model and their model is provided because their model is (to the best of our knowledge) the most recent reference on one-day-ahead electricity price forecasting apart from the most recent DFM we have already used as a benchmark.

After analysing the results shown in Table 2 in detail, it can be stated that:

- For six out of the nine months considered, our methodology is superior to all the models provided in Cruz *et al* (2011), although in principle the comparison is not fair since much more information is available in their case, when including the demand as an explanatory variable, and more parameters estimated.
- In the other three months, we still obtain better results than four out of their six models in February 2008 and five out of six models in January 2008 (note that we only provide in Table 2 their three best results among the six they provide in their paper).

[8]For the month of December 2007, the MAPE2, which is more robust to the presence of spikes in forecasting errors, is 9.5%, and the apparently low-quality result for this month is due to some particular spikes on Christmas Eve. Note that forecasts are computed for every day and hour in the period considered, even in the case in which the hourly price could be considered as an outlier.

Table 3 The global performance of the new model proposed

	Mon	*Tue*	*Wed*	*Thu*	*Frid*	*Sat*	*Sun*
Q_{25}	4.384	3.619	3.815	3.619	4.069	5.010	4.856
Q_{50}	5.627	5.187	5.438	5.431	5.553	6.055	6.333
Q_{75}	8.244	7.867	7.540	7.520	8.093	8.141	8.665

Percentiles 75, 50 (median) and 25, Q_{75}, Q_{50} and Q_{25}, respectively, of the daily MAPEs by day of the week (January 2007-December 2009)

- In December 2007 we beat three out of six models they provide, although these are not shown in Table 2 since only the three most accurate ones can be encountered there. Again it should be emphasised that the performance of our combined forecasts in this month is not bad, bearing in mind that we compute forecasts for every hour and day, and if there is any spike, the error that we obtain is larger (as when using any other forecasting method) and this affects the MAPE. However, the MAPE2 (calculated with as the average of the daily medians instead of means) is 9.5%.
- To summarise, considering that they provide forecasts with six models, we obtained better forecasts in the 88.89% of the cases considered.

In addition to showing our model's superior performance in comparison with previous works, in Table 3, we provide general numerical results for the three years under study (2007–2009). Forecasts have been computed for each hour, and the prediction errors for the whole period are also shown, which are even more important than specific results for a few days or weeks because they give a general idea of the performance of the model over a large span of time.

Finally, given the computational effort made, we provide some details about the computation time, just to illustrate that although many models are estimated, the procedure is feasible from the applied perspective to compute one-day-ahead forecasts for real applications. The computational experiments have been performed using a Windows-based personal computer with a 64-bit eight-core i7 processor at 1.73 GHz and 8 Gb of RAM. Several computational metrics are presented below:

- The average required CPU time for forecasting 1 h using the UMM is 1.23 s, and for forecasting 24 h in advance with the DFM method we needed 23.5 and 24.4 s for two and three factors, respectively. The six univariate methods provided and the 12 multivariate ones are run in parallel, and hence the CPU times needed are $1.23 + 23.5 + 24.4$ s.
- The CPU time required to solve the continuous and linear optimisation problem (5) for calculating the weights to build the combined forecasts for each hour is approximately 0.05 s.

7. Conclusions

In this paper, a novel methodological approach that uses both time series models and nonlinear optimisation techniques is proposed to calculate one-day-ahead forecasts of electricity prices by means of combined forecasts that account for renewable energies.

The main idea is that several univariate and multivariate time series models are estimated and used to compute out-of-sample hourly forecasts for a large span of days (all those days were encountered in the years 2007, 2008 and 2009). Some of our models consider the available information about renewable energies, particularly wind and hydroelectric energy, and others do not.

An extensive computational experiment has been conducted, and by means of a DOE all the forecasting methods that result from considering different models and explanatory variables are compared. Then, in the second stage the best models in terms of forecasting accuracy according to the DOE, that is, those models that are statistically better, are combined by minimising the MAPE and solving several non-linear mathematical programming problems. This procedure allows us to compute the most adequate weights that minimise the MAPE when we combine forecasts.

The designed model is robust, as the effect of the day is eliminated in the DOE by including it as a block. In fact, the results do not change significantly when we consider time-varying weights and when we consider the weights to be constant over time. Thus, once the model has been designed, only the weights are updated daily when new data are available. This is a feasible task from the computational point of view, as solving each of the 24 nonlinear optimisation problems to obtain the weights used to combine the 18 forecasting methods for each hour of the forthcoming day takes approximately 0.05 s.

The numerical results are shown for the Iberian Market, but of course they could be applicable to any other liberalised power market.

An important conclusion that can be drawn is that the high penetration of wind energy in the Iberian Market in the last few years has made hydro power less important for forecasting the price of electricity.

Furthermore, the results are shown in terms of their forecasting accuracies not just for a few weeks, but for the large span of days that we considered (three years). Moreover, in those months for which other authors had computed forecasts with different methodologies in previous works, our novel proposal provided superior results in terms of its forecasting accuracy. This result held even in the case of models that considered more explanatory variables (Cruz *et al*, 2011), and thus estimated more parameters than our models.

The MAPE for the whole period considered (the years 2007, 2008 and 2009) is 8.12%, and the MAPE2 (the average of the daily median errors for the whole period) is 5.98%. When we focus on the period from November 2007 to January 2009, the MAPE is 6.74%, which is lower than the result of 7.39% that was obtained by the DFM (García-Martos *et al*, 2012). All of the methodologies provided by Cruz *et al* (2011) are also improved by our approach in the vast majority of cases, as was shown in the previous section. This result occurred although these authors consider a larger number of explanatory variables, and particularly the inclusion of demand is considered by them and in our approach. Consequently, they estimate a larger number of parameters in their models.

References

Amjady N (2006). Day-ahead price forecasting of electricity markets by a new fuzzy neural network. *IEEE Transactions on Power Systems* **21**: 887–896.

Bates J and Granger C (1969). The combination of forecasts. *Operations Research Quarterly* **20**: 451–468.

Bunn DW (2004). *Modeling Prices in Competitive Electricity Markets*. Wiley: Chichester, UK.

Clemen RT (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting* **5**: 559–583.

Conejo AJ, Contreras J, Espnola R and Plazas MA (2005). Forecasting electricity prices for a day-ahead pool-based electric energy market. *International Journal of Forecasting* **21**(3): 435–462.

Conejo AJ, Carrión M, Nogales FJ and Morales JM (2010). Electricity pool prices: Long-term uncertainty characterization for futures-market trading and risk management. *Journal of the Operational Research Society* **61**: 235–245.

Conejo AJ, Nogales FJ and Arroyo JM (2002). Price-taker bidding strategy under price uncertainty. *IEEE Transactions on Power Systems* **17**: 1081–1088.

Conejo AJ, Nogales FJ, Arroyo JM and García-Bertrand R (2004). Risk constrained self scheduling of a thermal power producer. *IEEE Transactions on Power Systems* **19**: 1569–1574.

Contreras J, Espnola R, Nogales FJ and Conejo AJ (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems* **18**(3): 1014–1020.

Cottet R and Smith M (2003). Bayesian modeling and forecasting of intraday electricity load. *Journal of the American Statistical Association* **98**(464): 839–849.

Cruz A, Muñoz A, Zamora JL and Espínola R (2011). The effect of wind generation and weekday on Spanish electricity spot price forecasting. *Electric Power Systems Research* **81**: 1924–1935.

De Menezes M, Bunn DW and Taylor JW (2000). Review of practical guidelines for combining forecasts. *European Journal of Operational Research* **120**: 190–204.

Frestad D (2008). Common and unique factors influencing daily swap returns in the Nordic electricity market. *Energy Economics* **30**: 1081–1097.

García-Martos C, Rodríguez J and Sánchez MJ (2007). Mixed models for short-run forecasting of electricity prices: Application for the Spanish market. *IEEE Transactions on Power Systems* **2**(2): 544–552.

García-Martos C, Rodríguez J and Sánchez MJ (2011). Forecasting electricity prices and their volatilities using unobserved components. *Energy Economics* **33**(6): 1227–1239.

García-Martos C, Rodríguez J and Sánchez MJ (2012). Forecasting electricity prices by extracting dynamic common factors: Application to the Iberian market. *IET Generation, Transmission and Distribution* **6**(1): 11–20.

González C, Juan J, Mira J, Sánchez MJ and Prieto FJ (2005). Reliability analysis for systems with large hydro resources in a deregulated electric power market. *IEEE Transactions on Power Systems* **20**(1): 90–95.

Hippert HS, Pedreira CE and Souza RC (2001). Neural networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems* **16**(1): 44–55.

Huisman R, Huurman C and Mahieu R (2007). Hourly electricity prices in day-ahead markets. *Energy Economics* **29**: 240–248.

Jonsson T, Pinson P and Madsen H (2010). On the market impact of wind energy forecasts. *Energy Economics* **32**(2): 313–20.

Koopman SJ, Ooms M and Carnero MA (2007). Periodic seasonal reg-ARFIMA-GARCH models for daily electricity spot prices. *Journal of the American Satistical Association* **102**(477): 16–27.

Lam KF, Mui HW and Yuen HK (2001). A note on minimizing absolute percentage error in combined forecasts. *Computers & Operations Research* **28**: 1141–1147.

Lee RD and Carter LR. (1992). Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association* **87**(419): 659–671.

Misiorek A, Trück S and Weron R (2006). Point and interval forecasting of spot electricity prices: Linear *versus* non-linear time series models. *Studies in Nonlinear Dynamics and Econometrics* **10**(3): Article 2.

Montgomery DC (1984). *Design and Analysis of Experiments*. Wiley: New York.

Nogales FJ and Conejo AJ (2006). Electricity price forecasting through transfer function models. *Journal of the Operational Research Society* **57**: 350–356.

Nogales FJ, Contreras J, Conejo AJ and Espínola R (2002). Forecasting next-day electricity prices by time series models. *IEEE Transactions on Power Systems* **17**(2): 342–348.

Panagiotelis A and Smith M (2008). Bayesian density forecasting of intraday electricity prices using multivariate skewed *t* distributions. *International Journal of Forecasting* **24**: 710–727.

Peña D and Box GEP (1987). Identifying a simplifying structure in time series. *Journal of the American Statistical Association* **82**(339): 836–843.

Sánchez I (2006). Recursive estimation of dynamic models using Cook's distance, with application to wind energy forecast. *Technometrics* **48**: 61–73.

Saini LM, Aggarwal SK and Kumar A (2010). Parameter optimisation using genetic algorithm for support vector machine-based price-forecasting model in National electricity market. *IET Generation, Transmission and Distribution* **4**(1): 36–49.

Schwarz G (1978). Estimating the dimension of a model. *The Annals of Statistics* **6**(2): 461–464.

Shrivastava NA and Panigrahi BK (2011). A hybrid swarm-machine intelligence approach for day ahead price forecasting. *International Journal of Innovative Computing and Applications* **3**(4): 177–187.

Stock JH and Watson M (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association* **97**: 1167–79.

Stock JH and Watson MW (2004). Combination forecasts of output growth in a seven country dataset. *Journal of Forecasting* **23**: 405–430.

Taylor JW (2003). Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of the Operational Research Society* **54**: 799–805.

Taylor JW and Majithia S (2000). Using combined forecasts with changing weights for electricity demand profiling. *Journal of the Operational Research Society* **51**: 72–82.

Taylor JW and McSharry P (2007). Short-term load forecasting methods: An evaluation based on European data. *IEEE Transactions on Power Systems* **22**: 2213–2219.

Troncoso Lora A, Riquelme Santos JM, Gómez Expósito A, Martínez Ramos JL and Riqueline Santos JC (2007). Electricity market price forecasting based on weighted nearest neighbors techniques. *IEEE Transactions on Power Systems* **22**: 1294–1301.

Türkay BE and Demren D (2011). Electrical load forecasting using support vector machines: A case study. *International Review of Electrical Engineering* **6**(5): 2411–2418.

Vehvilainen I and Pyykkonen T (2005). Stochastic factor model for electricity spot price: The case of the Nordic market. *Energy Economics* **27**: 351–367.
Weron R (2006). *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach.* Wiley: Chichester.

Appendix

Daily MAPE statistics

Table A1 provides the basic metrics (average and standard deviation) of the daily MAPE for methods: UMM (García-Martos *et al*, 2007), DFM (García-Martos *et al*, 2012), and the *Combined Forecasts* method proposed in this paper, for the months from November 2007 to December 2011.

From Table A1 it is observed that not only is the average daily MAPE smaller for the combined forecasts, but also the standard deviation. (Be aware that the results here are not in percentage as in Table 1 but expressed as a per unit basis, ie, 10% is 0.1 expressed as a per unit basis.)

Table A1 Statistics for the daily MAPE

	UMM		*DFM*		*Combined*	
Month	*Avg.*	*Std. Dev.*	*Avg.*	*Std. Dev.*	*Avg.*	*Std. Dev.*
11-2007	0.091	0.026	0.089	0.027	0.078	0.021
12-2007	0.113	0.069	0.110	0.069	0.112	0.065
1-2008	0.101	0.043	0.095	0.046	0.090	0.052
2-2008	0.076	0.037	0.077	0.042	0.070	0.030
3-2008	0.090	0.032	0.080	0.031	0.067	0.027
4-2008	0.069	0.034	0.064	0.035	0.058	0.031
5-2008	0.062	0.032	0.056	0.031	0.055	0.033
6-2008	0.059	0.023	0.055	0.020	0.049	0.018
7-2008	0.053	0.018	0.047	0.017	0.044	0.015
8-2008	0.058	0.026	0.055	0.024	0.050	0.022
9-2008	0.055	0.024	0.054	0.022	0.044	0.014
10-2008	0.069	0.051	0.068	0.042	0.058	0.033
11-2008	0.075	0.051	0.071	0.045	0.060	0.035
12-2008	0.120	0.081	0.112	0.081	0.105	0.065